Modelling a Circadian Clock with HYPE

Modelling a Circadian Clock with HYPE

Vashti Galpin
Laboratory for Foundations of Computer Scince
School of Informatics, University of Edinburgh
Vashti.Galpin@ed.ac.uk

HYPE is a process algebra for hybrid systems. These systems contain both continuous and discrete behaviour. Modelling using HYPE involves identifying the flows that affect each continuous variable in the systems. This paper considers how to model the circadian clock of the green alga *Ostreococcus tauri* using this approach.

1 Introduction

Circadian clocks are biochemical networks with oscillatory behaviour that is adapted to the diurnal cycle. They consist of a number of biochemical species whose concentrations vary continuously over time. The discrete aspect of a circadian clock comes from the day/night cycle and the fact that the change from light to dark, and from dark to light can be viewed (and often is under laboratory conditions) as a discrete change in the system. In the traditional ordinary differential equation (ODE) models, this is usually approximated by an smooth function [4].

In this paper, I consider how to apply a more recent modelling technique which was developed for systems with both discrete and continuous behaviour. The process algebra HYPE provides the advantages of a process algebras (a small, elegant language with formal semantics) together with the ability to describe the flows in a hybrid model from which the ODEs can be obtained. This leads to a more compositional approach than that taken in other process algebras for hybrid system where the ODEs are required to appear monolithically in the syntax of the model.

The rest of the paper has the following structure: first HYPE is introduced, and then the circardian clock of *Ostreococcus tauri* is described after which its HYPE model is developed using two different approaches. Finally some conclusions are drawn.

2 HYPE

This section gives an informal (and incomplete) introduction to well-defined HYPE models. For more complete details, refer to [2, 3]. Basic HYPE subcomponents have the form

$$S(\vec{Y}) \stackrel{def}{=} \sum_{j=1}^{n} \underline{\mathrm{a}}_j{:}(\iota, r_j, I_j(\vec{X}_j)).S(\vec{Y}) + \underline{\mathrm{init}}{:}(\iota, r, I(\vec{X})).S(\vec{Y})$$

where $n \geq 0$, $\underline{\mathrm{a}}_j \neq \underline{\mathrm{a}}_k$ for $j \neq k$ and $\underline{\mathrm{a}}_j \neq \underline{\mathrm{init}}$ for all j. Each $\underline{\mathrm{a}}$ is an *event* and each $(\iota, r, I(\vec{Y}))$ is an *influence* or *activity* with ι an *influence name*, r an *influence strength* and $I(\vec{Y})$ an *influence type* which applies to a subset of the system variables, $\vec{Y} \subseteq \vec{X}$. Each subcomponent captures the discrete events and associated flows that affect a particular variable in the set of system variables. This link is expressed through the function iv which maps influence names to system variables. Each subcomponent contains only one influence name and an influence can only appear in one subcomponent. Moreover, each subcomponent must has an initial event $\underline{\mathrm{init}}$.

Submitted to:
PASTA 2010

Prefixes
$$\frac{}{\left\langle \underline{a} : (\iota, r, I).E, \sigma \right\rangle \xrightarrow{\underline{a}} \left\langle E, \sigma[\iota \mapsto (r,I)] \right\rangle} \qquad \frac{}{\left\langle \underline{a}.E, \sigma \right\rangle \xrightarrow{\underline{a}} \left\langle E, \sigma \right\rangle}$$

Cooperation
$$\frac{\left\langle E, \sigma \right\rangle \xrightarrow{\underline{a}} \left\langle E', \tau \right\rangle \quad \left\langle F, \sigma \right\rangle \xrightarrow{\underline{a}} \left\langle F', \tau' \right\rangle}{\left\langle E \underset{M}{\bowtie} F, \sigma \right\rangle \xrightarrow{\underline{a}} \left\langle E' \underset{M}{\bowtie} F', \Gamma(\sigma, \tau, \tau') \right\rangle} \quad \underline{a} \in M, \Gamma \text{ defined}$$

Figure 1: Selected rules for the operational semantics of HYPE

Subcomponents are assembled into components using $\underset{*}{\bowtie}$, and components into the uncontrolled system Σ. The symbol $\underset{*}{\bowtie}$ expresses the requirement that all shared events must be synchronised on.

The controlled system has the form $\Sigma \underset{*}{\bowtie} \underline{\text{init}}.Con$ where Con is a controller. Controllers have events but no influences, and are build up with the two-level syntax given by $M ::= \underline{a}.M \mid 0 \mid M + M$ and $Con ::= M \mid Con \underset{*}{\bowtie} Con$.

To complete the definition of a HYPE model, event conditions of the form $\text{ec}(\underline{a}) = (act(\underline{a}), res(\underline{a}))$ are required. For each event, $act(\underline{a})$ defines the conditions under which the event can occur in terms of Boolean formulae, and $res(\underline{a})$ defines the variables changes or resets that occur after an event. These are conjunctions of equations of the form $Y' = f(\vec{X})$ where the new value of the variable Y is defined as a function of the current values of the system variables.

Next to be considered is the operational semantics for a HYPE model. A *state* of the system is a function from influence names to pairs of influence strength and influence types, and a *configuration* consists of a controlled system together with a state $\left\langle ConSys, \sigma \right\rangle$. A state is a collection of flows rather than a valuation.

The operational semantics give a labelled transition system over configurations with events as labels. The rules are fairly standard, and the more interesting ones are given in Figure 1. For Prefix with influence, the state needs to be updated using $\sigma[\iota \mapsto (r,I)]$ which is defined by $\sigma[\iota \mapsto (r,I)](x) = (r,I)$ if $x = \iota$ and $\sigma[\iota \mapsto (r,I)](x) = \sigma(x)$ otherwise. For Cooperation over a shared action, the two new states in the premise of the rule need to be merged using the partial function Γ.

$$(\Gamma(\sigma, \tau, \tau'))(\iota) \quad = \quad \begin{cases} \tau(\iota) & \text{if } \sigma(\iota) = \tau'(\iota), \\ \tau'(\iota) & \text{if } \sigma(\iota) = \tau(\iota), \\ \text{undefined} & \text{otherwise.} \end{cases}$$

This function uses the previous state and the new states to determine which values have changed and then puts these changed values into the new state. Γ will be undefined if both the second and third argument differ from the first argument, namely if the values in the new state both differ from the old state since this represents conflicting updates.

The labelled transition system can then be used as a basis of a hybrid automaton to describe the hybrid behaviour of a HYPE model. Most importantly, each configuration $\left\langle CS, \sigma \right\rangle$ of the transition system is a mode in the hybrid automata and the ODEs defined for that mode are obtained from σ using the definition

$$CS_\sigma = \left\{ \frac{\mathrm{d}V}{\mathrm{d}t} = \sum \left\{ r \times [\![I(\vec{Y})]\!] \;\middle|\; \text{iv}(\iota) = V \text{ and } \sigma(\iota) = (r, I(\vec{Y})) \right\} \;\middle|\; V \in \vec{X} \right\}$$

where $[\![I(\vec{Y})]\!]$ is the meaning or interpretation of the influence type $I(\vec{Y})$. The rest of the mapping is straightforward and described in [3]. With the hybrid automata defined for a HYPE model, it is possible to obtain graphs describing the behaviour of the model over time.

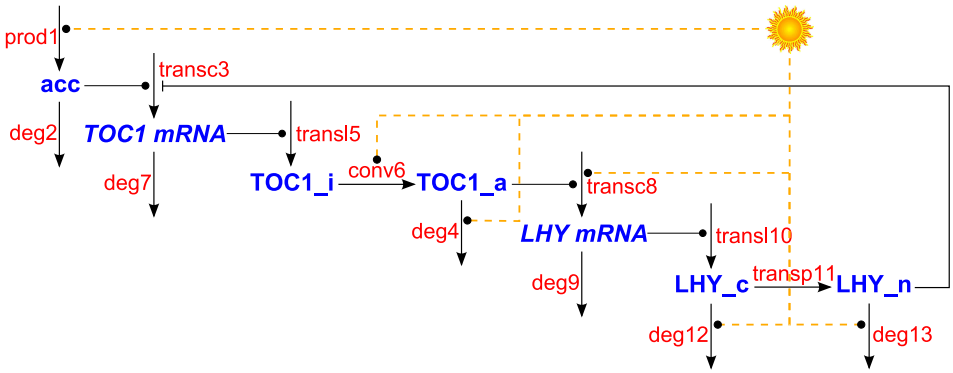

Figure 2: The genetic regulatory network of the Ostreococcus clock [4]

3 The circadian clock

Circadian clocks regulate various aspects of organisms to match the light/dark day/night cycle. They can entrain to the light conditions so that they are able to deal with changes in the length of light and dark periods. Typically, these clocks consist of a number of feedback loops. One of the simplest plant circardian clocks is that of the green alga *Ostreococcus tauri* [4, 1] and its hypothesised structure is described in Figure 2.

The functioning of this model can be described as follows. The transcription of TOC1 mRNA from the TOC1 gene is activated by light (represented here by a light accumulator *acc*) and inhibited by the LHY protein. TOC1 mRNA is translated to the inactive form of the protein TOC1 which becomes the active form of TOC1, more slowly in light than in dark. The degradation of the active form of TOC1 is enhanced by light. TOC1 activates the transcription of LHY mRNA from the LHY gene which is translated in LHY protein in the cytosol, and this translation is faster in light. It is transported to the nucleus where it inhibits transcription of the TOC1 gene. The degradation of LHY is enhanced by light.

This gives a time-based feedback loop where light and lack of LHY lead to production of TOC1 protein. When dark arrives, there is an increase in the amount of the active form of TOC1 which leads to more LHY being produced. The presence of LHY and dark lead to less TOC1, less activated TOC1 and hence less LHY. Once it is light again, the conditions are right for the production of TOC1 and the cycle repeats.

4 The HYPE model

Assuming just the information and diagram above, it is possible to start considering the HYPE model. It will have three events, init which is required, and light and dark with the following event conditions.

$$\begin{aligned} \mathrm{ec}(\underline{\mathrm{init}}) &= (\mathit{true}, (T' = t_0 \wedge (\mathit{initial\ values\ for\ all\ other\ variables})) \\ \mathrm{ec}(\underline{\mathrm{dark}}) &= (T = 12, \mathit{true}) \\ \mathrm{ec}(\underline{\mathrm{light}}) &= (T = 24, T' = 0) \end{aligned}$$

Hence dark happens after 12 hours and light after 24 hours, after which the time variable is reset. We refer to this patterns as LD 12:12. For reasons described below, we will assume the system starts in dark conditions by appropriate choice of t_0. The controller has a very simple form: $\mathit{Con} \stackrel{def}{=} \underline{\mathrm{dark}}.\underline{\mathrm{light}}.\mathit{Con}$.

From the diagram, the subcomponents for influences on a particular variable can be constructed. Apart from T for time, we have the variables A for *acc*, T_m for TOC mRNA, T_i for inactive TOC1 protein, T_a for active TOC1 protein, L_m for LHY mRNA, L_c for cytosolic LHY protein, and L_n for nucleic LHY protein. For each species variable, there will be an influence that increases the amount of the species and an influence that decreases it. Moreover, either of these two influences may differ depending on whether it is light or dark. Additionally, the strength of each influence must be considered and whether it is influenced by other variables, so that the influence type can be defined.

Subcomponents will have the form $I_{Y,i}(\vec{W})$ where Y is the species name, i the number of the reaction and $\vec{W}$ is a list of the variables which appear in the influence type. The reaction numbers are given in Figure 2 as the last part of the string associated with each reaction. All influence strengths, l_i, d_i and r_i are positive, hence the explicit use of the minus sign when they represent a negative influence strength. In the following, the constant influence type C has $[\![C]\!] = 1$, and the linear influence type $L(.)$ has $[\![L(Y)]\!] = Y$. We first need an influence to capture the passing of time since we have an explicit time variable T

$$Time \stackrel{def}{=} \underline{\text{init}}{:}(\iota_T, 1, C).Time$$

where $\text{iv}(\iota_T) = T$. We also define $\text{iv}(\iota_{Y,i}) = Y$.

The easiest influences to express are the degradation of species where the rate of degradation does not depend on light. It is known that the rate of degradation is linearly dependent on the quantity of species because the dynamics are mass action. So for the degradation of TOC1 mRNA, *acc*, TOC1 mRNA and LHY mRNA, the influences are

$$\begin{aligned}
I_{A,2}(A) &\stackrel{def}{=} \underline{\text{init}}{:}(\iota_{A,2}, -r_2, L(A)).I_{A,2}(A) \\
I_{T_m,7}(T_m) &\stackrel{def}{=} \underline{\text{init}}{:}(\iota_{T_m,7}, -r_7, L(T_m)).I_{T_m,7}(T_m) \\
I_{L_m,9}(L_m) &\stackrel{def}{=} \underline{\text{init}}{:}(\iota_{L_m,9}, -r_9, L(L_m)).I_{L_m,9}(L_m)
\end{aligned}$$

and these influences are always present. Next are the degradation influences that are affected by light, in that they are faster when there is light, hence in each case, $l_i > d_i$ for $i \in \{4, 12, 13\}$. Moreover, since the system starts in dark conditions $r_i = d_i$ and this is true for all influences where there are both r_i and d_i.

$$\begin{aligned}
I_{T_a,4}(T_a) &\stackrel{def}{=} \underline{\text{light}}{:}(\iota_{T_a,4}, -l_4, L(T_a)).I_{T_a,4}(T_a) + \underline{\text{dark}}{:}(\iota_{T_a,4}, -d_4, L(T_a)).I_{T_a,4}(T_a) + \\
&\quad \underline{\text{init}}{:}(\iota_{T_a,4}, -r_4, L(T_a)).I_{T_a,4}(T_a) \\
I_{L_c,12}(L_c) &\stackrel{def}{=} \underline{\text{light}}{:}(\iota_{L_c,12}, -l_{12}, L(L_c)).I_{L_c,12}(L_c) + \underline{\text{dark}}{:}(\iota_{L_c,12}, -d_{12}, L(L_c)).I_{L_c,12}(L_c) + \\
&\quad \underline{\text{init}}{:}(\iota_{L_c,12}, -r_{12}, L(L_c)).I_{L_c,12}(L_c) \\
I_{L_n,13}(L_n) &\stackrel{def}{=} \underline{\text{light}}{:}(\iota_{L_n,13}, -l_{13}, L(L_n)).I_{L_n,13}(L_n) + \underline{\text{dark}}{:}(\iota_{L_n,13}, -r_{13}, L(L_n)).I_{L_n,13}(L_n) + \\
&\quad \underline{\text{init}}{:}(\iota_{L_n,13}, -d_{13}, L(L_n)).I_{L_n,13}(L_n)
\end{aligned}$$

The production of *acc* is totally light dependent, hence the values for $r_1 = d_1$ are zero.

$$I_{A,1} \stackrel{def}{=} \underline{\text{light}}{:}(\iota_{A,1}, l_1, C).I_{A,1} + \underline{\text{dark}}{:}(\iota_{A,1}, 0, C).I_{A,1} + \underline{\text{init}}{:}(\iota_{A,1}, 0, C).I_{A,1}$$

The rates of translation of proteins from mRNA are linearly proportional to the amount of mRNA and not dependent on light conditions, so we have two production influences for the two types of mRNA in the model.

$$\begin{aligned}
I_{T_i,5}(T_m) &\stackrel{def}{=} \underline{\text{init}}{:}(\iota_{T_i,5}, r_5, L(T_m)).I_{T_i,5}(T_m) \\
I_{L_c,10}(L_m) &\stackrel{def}{=} \underline{\text{init}}{:}(\iota_{L_c,10}, r_{10}, L(L_m)).I_{L_c,10}(L_m)
\end{aligned}$$

There are two places in the model where we wish to model one species becoming another. First, the inactive form of the TOC1 protein becomes the active form, so there are two influences, one for the decrease in inactive TOC1 and one for the increase in active TOC1. This activation is faster in dark conditions, so $l_6 < d_6 = r_6$. This is modelled using mass action and due to the presence of a single species on the left-hand side of the chemical equation, there is a linear dependence.

$$\begin{aligned} I_{T_i,6}(T_i) &\stackrel{def}{=} \underline{\text{light}}{:}(\iota_{T_i,6}, -l_6, L(T_i)).I_{T_i,6}(T_i) + \underline{\text{dark}}{:}(\iota_{T_i,6}, -d_6, L(T_i)).I_{T_i,6}(T_i) + \\ &\quad \underline{\text{init}}{:}(\iota_{T_i,6}, -r_6, L(T_i)).I_{T_i,6}(T_i) \\ I_{T_a,6}(T_i) &\stackrel{def}{=} \underline{\text{light}}{:}(\iota_{T_a,6}, l_6, L(T_i)).I_{T_a,6}(T_i) + \underline{\text{dark}}{:}(\iota_{T_a,6}, d_6, L(T_i)).I_{T_a,6}(T_i) + \\ &\quad \underline{\text{init}}{:}(\iota_{T_a,6}, r_6, L(T_i)).I_{T_a,6}(T_i) \end{aligned}$$

The movement of the LHY protein from cytosol to nucleus can also be viewed as a change in species, and gives two influences, also with linear dependence.

$$\begin{aligned} I_{L_c,11}(L_c) &\stackrel{def}{=} \underline{\text{init}}{:}(\iota_{L_c,11}, -r_{11}, L(L_c)).I_{L_c,11}(L_c) \\ I_{L_n,11}(L_c) &\stackrel{def}{=} \underline{\text{init}}{:}(\iota_{L_n,11}, r_{11}, L(L_c)).I_{L_n,11}(L_c) \end{aligned}$$

The last two influences to consider are the transciption of mRNA from the genes. These cannot be expressed by a simple linear formula, and hence, unspecified functions with variables for the species involved are used.

$$\begin{aligned} I_{T_m,3}(A, L_n) &\stackrel{def}{=} \underline{\text{init}}{:}(\iota_{T_m,3}, 1, g(A, L_n)).I_{T_m,3}(A, L_n) \\ I_{L_m,8}(T_a) &\stackrel{def}{=} \underline{\text{light}}{:}(\iota_{L_m,8}, 1, f(T_a)).I_{L_m,8}(T_a) + \underline{\text{dark}}{:}(\iota_{L_m,8}, 1, f'(T_a)).I_{L_m,8}(T_a) + \\ &\quad \underline{\text{init}}{:}(\iota_{L_m,8}, 1, f'(T_a)).I_{L_m,8}(T_a) \end{aligned}$$

Now that all influences are defined, they can be put in cooperation to give the uncontrolled system Σ, which can then be put in cooperation with $\underline{\text{init}}.Con$ to give the full system.

4.1 Model experimentation and validation

To experiment with this model, parameters are needed for the influence strengths, initial values are required for all species variables, and definitions are required for the three functions f, f' and g. In the paper by Akman et al [1], there is a Bio-PEPA model which specifies various parameters which can be used in the HYPE model. This model assumes dark conditions at the start. The Bio-PEPA model also allows for the extraction of ODEs, and these have been validated against the original ODE model of the clock [4].

It is possible to do some validation of the HYPE model against the Bio-PEPA and its ODEs. Both the Bio-PEPA model and the HYPE model can be modified to execute in constant light conditions and in constant dark conditions. In the HYPE model to achieve constant light, this requires changing the influences following $\underline{\text{init}}$ to match those of the light and to change $\text{ec}(\underline{\text{dark}})$ to $(false, true)$ so that it never fires. A similar modification can be made to model constant dark. Graphs of the behaviour of some of the species are given in Figure 3 and they are indistinguishable from the graphs generated from the ODEs obtained from the Bio-PEPA model, which is to be expected since in these conditions the HYPE model is not hybrid. Figure 4 compares the graphs for the Bio-PEPA ODEs and the HYPE model under a 24-hour light/dark cycle with equal amounts of light and dark. The graphs are very similar; however, the peaks of the total LHY are slightly lower in the HYPE model. The graphs from the two models are also similar for the light conditions LD 6:18 and LD 18:6. This result is not unexpected as the Bio-PEPA model uses a switch function to model light and dark.

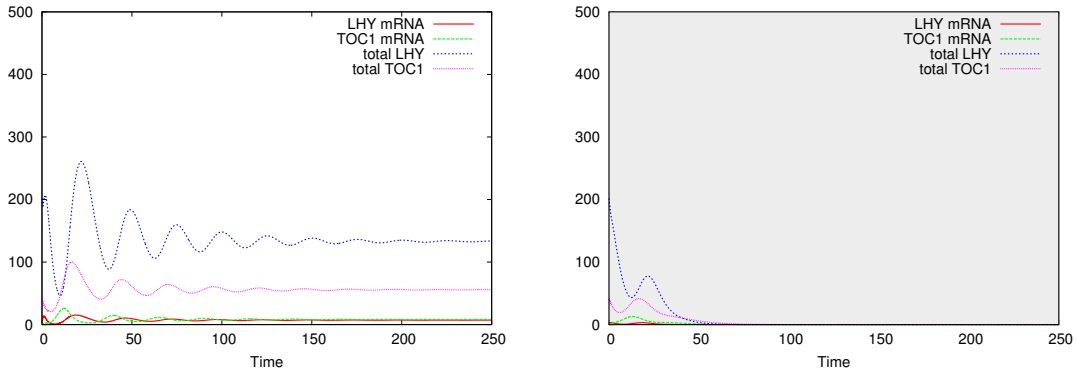

Figure 3: The HYPE model in constant light (left) and constant dark(right)

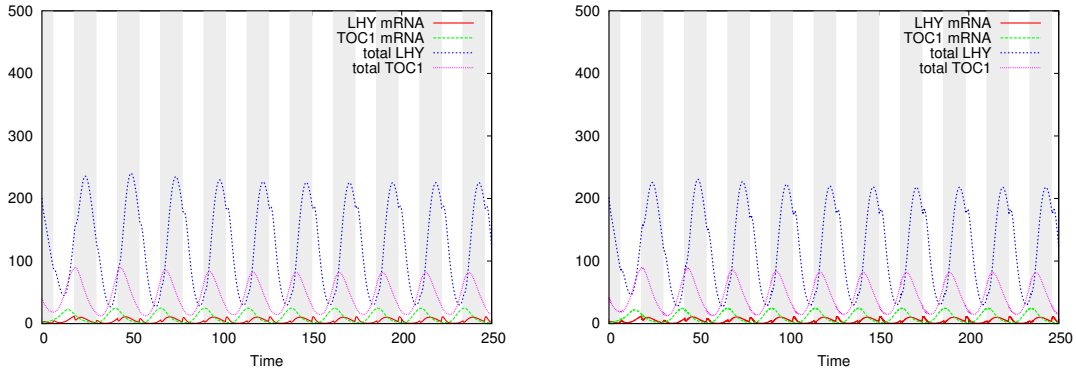

Figure 4: The Bio-PEPA ODEs (left) and the HYPE model (right) for LD 12:12

5 A different approach

As mentioned in the previous section, given a Bio-PEPA model, it is possible to extract ODEs from this model. If we have ODEs for each species then we can directly construct a hybrid model in the following manner. Assuming n species $S_1,\ldots,S_n$ and m reactions $R_1,\ldots,R_m$ involving only these species, let the $n \times m$ integer matrix D be the stoichiometry matrix for the species and reactions, and let the vector v be an $m \times 1$ matrix containing the vector laws for each reaction. In v, each element is a function over $\vec{X} = \{X_1,\ldots,X_n\}$, where X_i represents the amount or concentration of species S_i.

Then the ODEs for this system of species and reactions are $d\vec{X}/dt = D \times v$. To capture this in HYPE (without discrete events), we need to ensure for each ODE

$$\frac{dX_i}{dt} = \sum_{j=1}^{m} D[i,j]v[j] = \sum_{j=1}^{m} D[i,j]f_j(\vec{X})$$

that there is an influence for each $D[i,j]f_j(\vec{X})$. To this aim, define $[\![v[j]]\!] = f_j(\vec{X})$. Then define the following components.

$$I_{i,j}(\vec{X}) \stackrel{def}{=} \underline{\text{init}} : \alpha_{i,j}.I_{i,j}(\vec{X}) \quad \text{where} \quad \alpha_{i,j} = (\iota_{i,j}, D[i,j], v[j]) \quad \text{and} \quad \text{iv}(\iota_{i,j}) = X_i$$

Let $\mathrm{ec}(\underline{\mathrm{init}}) = (true, (X_1 = u_1) \wedge \ldots \wedge (X_n = u_n))$ where $u_1, \ldots, u_n$ are the initial values for the species. Then the overall HYPE model is

$$\big(\bowtie_{*}^{n}{}_{i=1} (\bowtie_{*}^{m}{}_{j=1} I_{i,j}(\vec{X})) \big) \bowtie_{*} \underline{\mathrm{init}}.0$$

In this formalisation, when $D[i, j] = 0$, subcomponents are defined whose influences have no effect. When considering the circadian clock, we omit these subcomponents.

To add discrete events, it is necessary to define the events in terms of what triggers them, and add them to the appropriate subcomponents with influences. Moreover, if events occur in specific sequences, then it is necessary to define a controller to specify this. Additionally, when events are added, it should be determined what the initial influence should be in each subcomponents where an event has been added.

If we apply this approach to the ODEs obtained from the Bio-PEPA model then we first need to establish what the ODEs are in the presence of constant light. This is straightforward to do as mentioned in the previous section. The HYPE model can then be constructed, with only $\underline{\mathrm{init}}$ as the discrete event. By understanding how the ODEs change when it is dark, it is then possible to add discrete events for $\underline{\mathrm{light}}$ and $\underline{\mathrm{dark}}$, and the appropriate influences for light-sensitive species. Finally, since it is assumed that the system starts in the dark, the initial influences must be changed in light-sensitive species.

Unsurprisingly, this gives a very similar model to that contructed in Section 4. The major differences is that all subcomponents are parameterised over all variables in the model constructed from the ODEs whereas in the model constructed from the diagrammactic representation, only the variables specifically required are used as parameters.

6 Conclusion

In this paper, a model of the *Ostreococcus tauri* circadian clock has been constructed in the process algebra HYPE based on a diagrammatic representation of the clock from which the influences on each species can be identified. Additionally, it has been shown that a similar model can be constructed directly from the ODEs of a Bio-PEPA model of the clock.

The modelling of the clock illustrates the approach that can be used in HYPE to find the flows in a model and construct subcomponents to represent these flows. Then from the HYPE model via a labelled transition system, a hybrid automaton can be constructed to describe the hybrid behaviour of the system over time. In this automaton, the ODEs at each node are constructed from the influences in the HYPE model, and this demonstrates the compositionality of the approach.

Acknowledgements: Thanks to Jane Hillston and Maria Luisa Guerriero for their comments. The author is supported by the EPSRC SIGNAL Project, Grant EP/E031439/1.

References

[1] O.E. Akman, M.L. Guerriero, L. Loewe & C. Troein (2010): *Complementary approaches to understanding the plant circadian clock*. In: *Proceedings of FBTC'10*, EPTCS, Vol. 19, pp. 1–19.

[2] V. Galpin, L. Bortolussi & J. Hillston (2009): *HYPE: a process algebra for compositional flows and emergent behaviour*. In: M. Bravetti & G. Zavattaro, editors: *Proceedings of CONCUR 2009*, LNCS 5710, pp. 305–320.

[3] V. Galpin, L. Bortolussi & J. Hillston (2010). *HYPE: hybrid modelling by composition of flows*. In preparation.

[4] J.S. O'Neill, G. van Ooijen, C. Troein, L. Dixon, F.-Y. Bouget, F. Corellou & A. Millar (2009). *Circadian rhythms persisting in the absence of transcription in a eukaryote*. Submitted for publication.